How to evolve the current AllEvents Micro sample is the most controversial part of this proposal, and the full CMWG2 committee must be involved in discussing these requirements. With requirement 6.1, the minimal (1Kbyte/event) AllEvents New Micro sample would be intended primarily only to start a new analysis, to select a desired subset of events from which to create a skim. It could also be used as the 'core' of a pointer-based skim, together with additional information per event and candidate. Publication-quality analysis would be performed from skims. Without requirement 6.1 we expect the AllEvents New Micro sample to be equivalent to the existing Micro in size and content, and support full analysis. To help guide the discussion, the sub-committee has prepared some arguments 'for' and 'against' requirement 6.1, which are presented in a separate document.

<center>Arguments against requirement 6.1</center>

1. It is what we have had up until now, and it will reduce the transition pains -- especially if it can be made so that a piece of user code which runs on the present micro also runs on the new micro ("as is" or with minimal modifications).
2. It gives more flexibility for analyses that do not fit in an existing skim -- especially if the requirement of "a new skim every three months" proves to be too ambitious.
3. For those skims that do not need the mini information to be made, it allows for faster recovery from a flawed skim algorithm, by running on the AllEvents sample or by reskimming from it.
4. The cost in terms of additional storage may be quite small. Take the following size estimates for micro:
    4.1. AllEvents, no quals ~ 1 Kb/event + 0.7 Kb/event (MC truth)
    4.2. AllEvents, with quals ~ 2 Kb/event + 0.7 Kb/event (MC truth
    4.3. Skims ~ 1.2 Kb/event (1 Kb no quals + 0.2 Kb custom) + 0.7 (MC truth)
5. Then, if we assume an event replication factor of 4 for the skims, the storage requirements in the two scenarios are
    5.1. Data:
        5.1.1. no quals: 1 Kb/event + 4*1.2 Kb/event = 5.8 Kb/ev
        5.1.2. with quals: 2 Kb/event + 4*1.2 Kb/event = 6.8 Kb/ev (17% more)
    5.2. MC:
        5.2.1. no quals: 1.7 Kb/event + 4*1.9 Kb/event = 9.3 Kb/ev
        5.2.2. with quals: 2.7 Kb/event + 4*1.9 Kb/event = 10.3 Kb/ev (11% more)
Of course the 17% (11%) increases if the skims make heavy use of pointers, which have advantages in terms of storage requirements and disadvantages in terms of complexity and access time.

Arguments for requirement 6.1

1. If the AllEvents New Micro sample contains all the detector information, it will be explicitly redundant with the skims and (to some extent) the Mini, which is not an optimal use of resources.
2. Running analysis on the AllEvents sample is inefficient compared to running the same analysis on a skim.
3. Most analyses which have been run on AllEvents could be migrated to run on an appropriate skim without code changes.
4. The fact that the 10-series skims replicated the data 4 times is not a good measure to apply to the new model, as that will support both pointer and deep copy skims. I estimate that a factor of 2 replication is probably achievable.
5. It is clear that, when transitioning to the new Analysis Model, a copy of the AllEvents data with complete content will need to be made available. However, as BaBar adopts the new model by making better use of skims, and as our data sample grows, we should transition to keeping only a minimal version of the AllEvents sample to make optimal use of our resources.